BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :		(11) International Publication Number: WO 93/24631
C12N 15/31, 15/75, 15/62 C12N 15/54	A1	(43) International Publication Date: 9 December 1993 (09.12.93)

(21) International Application Number: PCT/US93/05240

(22) International Filing Date: 27 May 1993 (27.05.93)

(30) Priority data:
07/891,524 29 May 1992 (29.05.92) US

(71) Applicant: E.I. DU PONT DE NEMOURS AND COMPANY [US/US]; 1007 Market Street, Wilmington, DE 19898 (US).

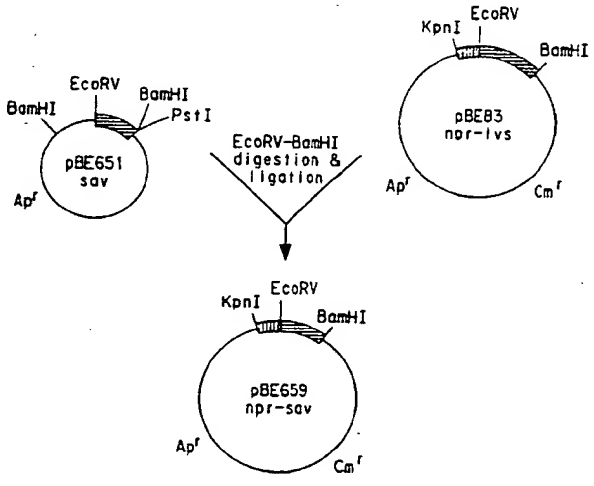
(72) Inventor: NAGARAJAN, Vasantha ; 13 Dickinson Lane, Wilmington, DE 19807 (US).

(74) Agents: GEIGER, Kathleen, W. et al.; E.I. du Pont de Nemours and Company, Legal/Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).

(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PRODUCTION OF STREPTAVIDIN FROM *BACILLUS SUBTILIS*



(57) Abstract

Method for producing biologically active streptavidin and streptavidin fusion proteins by cloning the *sav* gene from *Streptomyces avidinii* into *Bacillus subtilis* and purifying the secreted streptavidin proteins from the growth medium.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**MR**	Mauritania
AU	Australia	**GA**	Gabon	**MW**	Malawi
BB	Barbados	**GB**	United Kingdom	**NL**	Netherlands
BE	Belgium	**GN**	Guinea	**NO**	Norway
BF	Burkina Faso	**GR**	Greece	**NZ**	New Zealand
BG	Bulgaria	**HU**	Hungary	**PL**	Poland
BJ	Benin	**IE**	Ireland	**PT**	Portugal
BR	Brazil	**IT**	Italy	**RO**	Romania
CA	Canada	**JP**	Japan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SK**	Slovak Republic
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SN**	Senegal
CM	Cameroon	**LI**	Liechtenstein	**SU**	Soviet Union
CS	Czechoslovakia	**LK**	Sri Lanka	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**MC**	Monaco	**UA**	Ukraine
DK	Denmark	**MG**	Madagascar	**US**	United States of America
ES	Spain	**ML**	Mali	**VN**	Viet Nam
FI	Finland	**MN**	Mongolia		

TITLE

PRODUCTION OF STREPTAVIDIN FROM *BACILLUS SUBTILIS*

FIELD OF THE INVENTION

This invention relates to a method for cloning the streptavidin gene into *B. subtilis* and the secretion of tetrameric, biologically active streptavidin protein into the growth medium.

BACKGROUND OF THE INVENTION

Streptavidin is a tetrameric protein isolated from the actinobacterium *Streptomyces avidinii* and is remarkable for its ability to bind up to four molecules of d-biotin with unusually high affinity (1). Streptavidin is a nearly neutral 60,000 dalton protein consisting of four identical subunits each having a molecular weight of 15,000 daltons (2). The ability of streptavidin to bind derivitized forms of biotin has led to its widespread use in diagnostic assays where high affinity protein-ligand interactions are important. Some of the current applications include streptavidin coated liposomes used for drug delivery and diagnostic tests to detect human antibodies or pathogens using streptavidin linked to enzymes such as alkaline or phosphatase or horseradish peroxidase. Streptavidin is currently produced in commercial quantities by *S. avidinii*. *S. avidinii* naturally secretes relatively low amounts of streptavidin into the growth media. Recombinant *E. coli* produce streptavidin in relatively high amounts, but only intercellularly and generally in insoluble form. Streptavidin is available commercially from several manufacturers, however, the commercial reagent exhibits considerable variation in molecular weight and purity between suppliers and between lots from the same supplier. Furthermore, the reagent is expensive and its cost prohibits broader applications.

The primary cause for the lack of homogeneity in the streptavidin currently produced is the presence of protease susceptible sites in the protein which exist outside of the biotin binding domain. Another factor contributing to the low purity of streptavidin is the presence of trace amounts of biotin in the growth medium from which the streptavidin is purified. This results in biotin-bound streptavidin and contributes to the molecular weight variations mentioned above.

Argarana and Meade (2) and (3) describe the cloning of the streptavidin gene from a genomic library of *Streptomyces avidinii* as well as the DNA sequence of the coding region of the gene.

Meade also reports secretion of 250 mg/liter of streptavidin from *S. lividans*. This process is time-consuming as the fermentation time alone is 4 to 7 days.

Cantor (4) describes the isolation of the DNA which encodes streptavidin from *Streptomyces avidniii*, which includes the region encoding the signal peptide and the subsequent cloning of the DNA into a bacterial host cell, typically *E. coli*. Additionally, Cantor describes the construction and subsequent expression in bacteria of a fused gene comprising a first DNA fragment encoding a target protein of interest (specifically human LDL receptor) fused to a DNA fragment encoding streptavidin.

Sano and Cantor (5) describe the construction of systems for expressing the cloned streptavidin gene in *E. coli* where the streptavidin accumulated to more than 35% of the total cell protein. Sano further describes the creation of expression vectors for streptavidin containing chimeric proteins which are also capable of expression in *E. coli* (6). Meade shows that streptavidin is present in the periplasm of *E. coli*. However, the work of Sano and Cantor has indicated that

the str ptavidin signal peptide does not function in *E. coli*.

The methods of Sano and Cantor are useful and provide an alternative to the production of streptavidin from *Streptomyces avidinii*. However, a significant problem with the production of streptavidin from *E. coli* is that the protein is expressed intracellularly and in insoluble form. Recovery of useful protein from this system generally requires methods for isolation of the inactive streptavidin followed by purification and refolding of the protein into the active form. These methods are time-consuming and not readily adapted for commercial production since refolding of the protein is not efficient.

B. subtilis, a gram positive bacterium, has great potential for producing commercially important proteins because it can be genetically manipulated, adapted to various nutritional and physical conditions of growth, and because it is not pathogenic or toxigenic to humans. Under the proper conditions, *B. subtilis* is known to synthesize and secrete specific proteins relatively free of contaminating species making the proteins easier to purify.

Chang, Nagarajan and Koracevic (7), (8), and (9) separately disclose cloning vectors capable of replication in *B. subtilis* which include heterologous genes coding for proteins which can be secreted into the bacterial growth media. Proteins that may be secreted by these systems include proteases, Protein A, prorennin, insulin, human growth hormone, interferon and urokinase.

It is generally understood that translocation of secreted proteins across bacterial membranes requires a signal peptide. While a number of studies directed to understanding the role of the signal peptide in protein

secretion have been done, the mechanism of such translocation and the exact manner by which the signal peptide influences translocation and removal of the signal peptide from the signal peptide-mature protein complex to yield secreted mature protein is not fully understood.

Vectors enabling the secretion of a number of different heterologous proteins by *B. subtilis* have been demonstrated (8), (9), and (10). These include vectors that are based on genes for bacterial exoenzymes such as amylase, protease, levansucrase and β-lactamases.

Palva demonstrated the secretion by *B. subtilis* of the heterologous proteins β-lactamase of *E. coli* (12) and human leukocyte interferon (13) by transforming the bacteria with a vector wherein the genes for β-lactamase of *E. coli* and human leukocyte interferon were operably linked to the promoter, ribosome binding site and signal sequence of the α-amylase gene from *Bacillus amyloliquefaciens*. They found that only a low amount of interferon was secreted.

Secretion vectors based on levansucrase were reported by Dion (14). Dion obtained low levels of mouse interferon compared to levansucrase. The authors suggested that the low yield of interferon was due to a poorly understood incompatibility between the signal sequence of the *B. subtilis* levansucrase gene and the mouse interferon α2 gene.

Nagarajan (15) describes a method to design vectors for the secretion of heterologous proteins in bacteria, including *B. subtilis*. The method is drawn to ways of enabling combinations of promoters, ribosome binding sequences and signal peptides with sequences from desired heterologous proteins such that translocation via an effective signal peptide is achieved. However, in the case of streptavidin secretion by *B. subtilis*,

the method disclosed will not ensure a commercially viable secretion process. It is known, for example, that the tetrameric form of streptavidin protein is needed for efficient biotin binding and prior to the instant invention it had not been demonstrated that *B. subtilis* could efficiently secrete and accumulate a tetrameric protein. Most of the proteins that have been efficiently secreted to date are monomeric proteins, with the exception of *E. coli* alkaline phosphatase which is dimeric. Also, the production of biologically active streptavidin in the *B. subtilis* growth medium requires several processes to occur simultaneously and efficiently. These include the efficient translocation of the mature protein across the membrane and removal of signal peptide from the precursor protein. This must be followed by the release of the mature protein from the membrane and oligomerization of the monomers to yield a tetrameric protein which passes into the medium. It is not clear from current knowledge whether the oligomerization in *B. subtilis* occurs in the space between the cell membrane and cell wall, or if it occurs in the growth medium. Further, it is known that the *Streptomyces* genome contains a high GC base pair content of at least 70%, whereas the GC content in *B. subtilis* is known to be on the order of 42%. This would indicate that translation of a *Streptomyces* gene by *B. subtilis* may not be possible due to an incongruous number of GC codons in the *Streptomyces* genes. Furthermore, it is well known that *B. subtilis* secretes some proteins very inefficiently, which, in the case of streptavidin, would prove lethal to the cell.

Therefore, a need remains for an expression system that is capable of expressing and secreting streptavidin protein as an active, soluble protein in commercially useful quantities which eliminates the need for

additional solubilization and purification. The instant invention provides a method for efficiently expressing and secreting soluble, active streptavidin from *B. subtilis* in high quantities virtually free of biotin and other contaminants.

SUMMARY OF THE INVENTION

This invention provides a method for producing tetrameric, biologically active streptavidin by secretion from *Bacillus subtilis*, comprising:

a. transforming *Bacillus subtilis* with a gene construct comprising a sequence encoding streptavidin operably linked to a sequence encoding a signal peptide and an expression element wherein said sequence encoding a signal peptide is isolated from DNA encoding exproteins of bacteria, and said expression element is isolated from DNA encoding gram positive bacterial proteins;

b. growing the transformed *Bacillus subtilis* in suitable growth medium whereby streptavidin is secreted into the growth medium; and

c. purifying the streptavidin from the growth medium.

This invention further provides a *Bacillus subtilis* bacterium transformed as described in (a) above.

This invention also provides a method for producing a fused gene product comprised of tetrameric, biologically active streptavidin fused to a second desired protein, by secretion from *Bacillus subtilis*, comprising:

a. transforming *Bacillus subtilis* with a fused gene construct comprising a sequence encoding the streptavidin gene fused to a sequence encoding a second desired protein wherein said fused sequence is operably linked to a sequence encoding a signal peptide and an expression element; wherein said sequence encoding a

signal peptide is isolated from DNA encoding exoproteins of bacteria and said expression element is isolated from DNA encoding gram positive bacterial proteins;

b. growing the transformed *Bacillus subtilis* in suitable growth medium whereby the fused streptavidin and desired protein is secreted into the growth medium; and

c. purifying the fused streptavidin and desired protein from the growth medium.

This invention further provides a *Bacillus subtilis* bacterium transformed with the fused gene construct as described in (a) above, and also the fused gene products as described above.

Also, plasmids pBE659, pBE660, pBE661, pBE662, pBE663, pBE673 and pBE655 are provided.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1a describes the construction of a plasmid pBE659 containing a hybrid gene fusion construct consisting of *sav* gene fused to *npr* expression elements and signal peptide from the plasmids pBE651(*sav*) and pBE83(*npr-lvs*). A) The DNA sequence across *npr* signal peptide cleavage site in pBE83 is designated as SEQ ID NO:10. B) The DNA sequence across *sav* mature protein sequence in pBE651 is designated as SEQ ID NO:11. C) The DNA sequence across *npr-sav* fusion junction in pBE659 is designated as SEQ ID NO:12.

Figure 1b describes the DNA sequence of the streptavidin gene. The DNA sequence is designated as SEQ ID NO:13. The amino acid sequence is designated as SEQ ID NO:14.

Figure 2 describes the construction of plasmids containing the gene fusion constructs pBE660(*apr-sav*), pBE661(*npr-sav*), pBE662(*bar-sav*), and pBE663(*lvs-sav*) from the plasmids pBE30(*apr-phoA*), pBE90(*npr-phoA*), pBE91(*bar-phoA*), and pBE597(*lvs-phoA*), respectively.

Figure 3 is a Western Blot analysis of *B. subtilis*, pBE20, and pBE659 culture supernatant, respectively. Numbers denote time in hours.

Figure 4a depicts the U.V. absorption spectra of fractions eluted from an iminobiotin agarose column loaded with an ammonium sulfate fraction of *B. subtilis* growth media.

Figure 4b depicts the U.V. absorption spectra of fraction eluted from a Sephacryl S200 column loaded with the streptavidin-containing fraction from the iminobiotin column.

Figure 4c is a Western Blot analysis of fractions from both the iminobiotin and S200 columns.

Figure 5a describes the creation of plasmid pBE93(*npr-phoA*) from plasmid pBE592(*lvs-phoA*) by restriction enzyme digestion and the ligation of the *npr* expression element DNA from pBE93 with the mature *sav* gene of plasmid pBE670 by Nhe1-Pst1 digest resulting in the creation of plasmid pBE673(*npr-sav*).

Figure 5b depicts a Western blot of *B. subtilus* strains containing pBE659(1-159) and pBE673(15-159).

Figure 6 describes the streptavidin-heterologous gene (PhoA) fusions: Npr_{ss}-Sav_{15-133}-PhoA; Npr_{ss}-Sav_{15-138}-PhOA; Npr_{ss}-Sav_{15-145}-PhoA; and Npr_{ss}-Sav_{15-159}-PhoA.

Figure 7 describes the EcoRV digestion and ligation of the *lvs* expression element and mature *lvs* gene from plasmid pBE311 with the mature *sav* gene of pBE653(*sav*) to produce the plasmid pBE655 containing the gene fusion *lvs-sav-lvs*.

DETAILED DESCRIPTION OF THE INVENTION

In support of the disclosure of the instant invention, the following terms are intended to convey the following meanings.

"Mature protein" is the final protein product without the signal peptide attached.

"Desired protein" is any protein considered a valuable product to be obtained from genetically engineered bacteria. The term "desired protein," or "second desired protein" is used herein to describe that protein which, in the fused gene product of the invention, is fused to streptavidin. Also, unless otherwise indicated, Applicants intend, according to the instant invention, that the fused gene construct may be constructed so that the streptavidin is fused to either the C-terminus, or the N-terminus, of the second desired protein.

"Signal peptide" is an amino terminal polypeptide preceding the secreted mature protein. The signal peptide is cleaved from and is therefore not present in the mature protein. Signal peptides function by directing and translocating extracellular proteins across cell membranes. Signal peptide is also referred to as signal protein.

"Compatible restriction sites" are different restriction sites which, when cleaved, yield nucleotide ends that can be ligated without any additional modification.

"*apr*" and "Apr" refer to alkaline protease gene and protein, respectively.

"*bar*" and "Bar" refer to ribonuclease gene and protein, respectively.

"*lvs*" and "Lvs" refer to levansucrase gene and protein, respectively.

"*npr*" and "Npr" refer to neutral protease gene and protein, respectively.

"*phoA*" and "PhoA" refer to *E. coli* alkaline phosphatase gene and protein, respectively.

"*sav*" and "Sav" refer to streptavidin gene and protein, respectively.

The term "streptavidin" refers to the protein comprising amino acid residues 1-159 of a 2 kb BamHI fragment isolated from *streptomyces avidinii* as described in Example 1 and Figure 1b, or any sequential subset of amino acid residues thereof, or mutations or derivatives thereof, which retain the ability to bind biotin.

The term "biotin" refers to biotin, biotin derivatives, and biotin analogs capable of binding streptavidin.

"Ap-" refers to ampicillin.

"Kan-" refers to kanamycin.

"Cm-" refers to chloramphenicol.

"Shuttle phagemid" as used herein is a vector that is normally double stranded and contains both the origins of replication for *E. coli* and *B. subtilis* and also the F1 intragenic region for the preparation of single stranded DNA.

The terms "peptide", "polypeptide" and "protein" are used interchangeably.

The terms "restriction endonuclease cleavage site" and "restriction site" are used interchangeably.

The term "expression element" refers herein to a DNA fragment containing the necessary information for transcription and translational initiation in *B. subtilis* including, for example, the promoter sequence and the ribosome binding site sequence.

The term "exoproteins of bacteria" is used to describe those proteins produced by bacteria which are able to cross the cytoplasmic bacterial membrane and are known to be naturally secreted into the bacterial growth media.

The term "gram positive bacterial proteins" refers to those proteins which are known to be naturally synthesized by gram positive bacteria.

The term "biologically active", as used herein, refers to a soluble streptavidin protein molecule which is able to bind to biotin, biotin analogs, or derivatized biotin.

Suitable methods of genetic engineering employed herein are described in the references (16), (17), and (18) and in the instructions accompanying commercially available kits for genetic engineering. Bacterial cultures and plasmids necessary to carry out this invention are commercially available and, along with their sources, are identified in the text and examples which follow.

The source of the bacteria strains, genes and the various vectors described herein are readily available to one skilled in the art. The complete nucleotide sequence for the genes for *apr*, *npr*, *bar* and *lvs* from *B. amyloliquefaciens*, the *E. coli phoA* gene and the *sav* gene from *S. avidinii* have been published (2), (19), (20), (21), and (22). In addition, these sequences are accessible in the GenBank from Nucleic Acid Data base from Los Almos, California.

The bacterial strains *B. subtilis*, *E. coli* and *S. avidinii* and plasmids pBE322, pTZ18R, pSK, pC194, pUB110 and phage M13KO7 are readily available from a variety of sources. For example, they can be obtained from the American Type Culture Collection, Rockville, MD, or the Bacillus Stock Center, Ohio, and are also available from other commercial suppliers. The position of the newly engineered restriction sites and sequence of the mutagenic oligonucleotide is indicated in the figures or examples and one skilled in the art may

readily prepare these constructs with the available information in this art.

The anti-streptavidin antiserum can be purchased from several manufacturers. In the present study, it was purchased from Sigma, St. Louis, MO. The streptavidin used as standard was obtained from Bethesda Research Laboratories, MD. The various techniques used for the construction of plasmids and bacterial strains are standard methods and relevant variations are indicated in the text when necessary. These methods are found in references (11), (16), (17), and (18).

The present invention utilizes an isolated 2 kb DNA fragment which encodes streptavidin (Figure 1b). The DNA was isolated according to techniques well known in the art based on published DNA sequence (2, 3). The 2 kb fragment contains the entire *sav* open reading frame encoding a signal peptide and the mature protein which comprises amino acid sequences 1-159 and the flanking region DNA which occurs naturally at the 3' and 5' ends of the coding region.

A recombinant cloning vehicle is described which comprises DNA encoding a suitable expression element for streptavidin expression and the DNA fragment encoding the streptavidin protein, wherein said cloning vehicle is further characterized by the presence of a first and a second restriction enzyme site, the DNA fragment encoding streptavidin being inserted into said site.

The present invention entails development of a vector which comprises expression elements including a promoter sequence controlling transcription and a ribosomal binding site sequence controlling translation; and also a sequence for a signal peptide which enables translocation of the protein through the bacterial membrane and the cleavage of the signal peptide from the mature protein. Suitable vectors will be those which

are compatible with the bacterium employed. For example, for *B. subtilis* such suitable vectors include *E. coli-B subtilis* shuttle vectors which have compatible regulatory sequences and origins of replication. They will be preferably multicopy and have a selective marker gene, for example, a gene coding for antibiotic resistance. For example, pTZ18R is a phagemid obtainable from Pharmacia, Piscataway, NJ 08854, which confers resistance to ampicillin in *E. coli*, and pC194 from BGSC which confers resistance to chloramphenicol (cm^r) in *E. coli* and *B. subtilis*.

The expression elements containing DNA sequences encoding the promoter and ribosome binding site may be from any gram positive bacterial protein, and the signal peptide may be from any single bacterial gene which encodes a secreted product. The DNA sequences encoding the promoter and ribosome binding site may also be from a different gene than that encoding the signal peptide. These DNA sequences encoding the promoter, ribosome binding site and signal peptide can be isolated by means well known to those skilled in the art and illustrative examples are documented in the literature (23). The promoters in the DNA sequences may be either constitutive or inducible. Suitable signal peptides and expression elements may be selected from the group comprising, for example, *apr*, *npr*, *lvs* and *bar*.

The addition of a restriction endonuclease cleavage site to the 3' end of the DNA encoding the signal peptide is also easily accomplished by means well known to those skilled in the art (16). Several methods may be employed to add the restriction endonuclease cleavage site to the 3' end of the DNA encoding the signal peptide. One such method might incorporate polymerse chain reaction (PCR). The method used in the present invention is site-directed mutagenesis which is most

preferred, and is described in Mutagene manual (Biorad, 1414 Harbour Way South, Richmond, CA 96804). The means to isolate DNA sequences encoding a desired protein and the addition of restriction sites on the 5' end of the mature coding sequence is also well known to those skilled in the art (16). Any restriction endonuclease site may be used but the use of a restriction site unique to the target vector is desirable. The restriction endonuclease site on the 3' end of the DNA sequence encoding the signal peptide and that on the 5' end of the DNA sequence encoding the mature desired protein must be compatible. Suitable compatible restriction sites are well known in the art. See, for example the Restriction Fragment Compatibility Table of the New England Biolabs 1988-1989 Catalog, New England Biolabs, Inc., Beverly, MA 01915 (1988), which is herein incorporated by reference. Preferred for use herein are *EcoRV* or *NheI*. The combined DNA sequences encoding a promoter, ribosome binding site and signal peptide with a restriction site at its 3' end and the DNA sequences encoding mature polypeptides or proteins with a compatible restriction site at its 5' end can be operably integrated by conventional techniques (16) and (17).

The recombinant cloning vehicle of the present invention has been inserted into a bacterial host cell. A suitable host cell would be derived from the genus *Bacillus*, the most preferred host cell would be of the species *subtilis*. One method to transform *B. subtilis* bacteria is described by Vasantha et al. (9). Standard microbiological methods well known to those skilled in the art can be used for the growth and maintenance of bacterial cultures. Several genetically engineered *B. subtilis* host cells containing the recombinant cloning vehicle of the present invention have been

prepared by transforming the strain BE1500 or its derivatives (1510) with th plasmids, pBE659A, pBE660, pBE661, pBE662, pBE663, pBE673, pBE659 and pBE655. BE1500 has the genotype trpC2, metB10, lys3, Δ-aprE66, Δ-npr82, Δ-sacB::ermC (24).

A method of producing streptavidin comprises cultivating a genetically engineered host cell of the present invention under suitable conditions permitting expression of the streptavidin gene and recovering the streptavidin so produced from the growth media.

The present invention also provides a fused gene which comprises a first DNA fragment encoding streptavidin fused to a second DNA fragment encoding a target protein of interest and wherein the fused gene is capable of expressing a fused protein in vivo when the gene is inserted into a host cell.

In one embodiment of the invention, the second DNA fragment is the gene encoding *B. amyloliquefaciens* levansucrase (Lvs). Such a fused gene expresses a protein which consists of streptavidin at the N-terminal region of the fused protein and levansucrase at the C-terminal region of the fused protein when the fused gene is inserted into a suitable expression vector and introduced into a suitable host cell. The fused gene may be cloned into a bacterial expression vector and used to transfect a bacterial host cell with the fused gene. A preferred bacterial host cell is *B. subtilis*.

The invention also provides a secretion vector capable of expressing and secreting the fused gene of the present invention when said vector is introduced into a suitable host cell. The vector comprises DNA encoding the promoter, ribosomal binding site and signal sequence of the first gene followed by DNA encoding the mature protein of the first gene, fused to DNA encoding the mature protein of the second gene.

Also provided is a fused protein encoded by the fused gene of the present invention wherein a desired protein of interest is fused to streptavidin. In one embodiment of the invention, the desired protein is *B. amyloliquefaciens* levansucrase.

For purification of the streptavidin protein product, bacteria secreting streptavidin are grown in a standard growth media such as S7 medium and are separated from the growth media by centrifugation. The proteins in the growth media may be concentrated either by membrane filtration techniques or ammonium sulfate precipitation or both, but most preferably by 70% ammonium sulfate precipitation. The concentrated proteins are reconstituted in an appropriate buffer compatible with binding to an iminobiotin affinity resin. The buffer for the reconstitution of the protein and the equilibration of the iminobiotin affinity resin is preferably about pH 8.5 to pH 11.0, but most preferably about pH 11.0. The concentrated protein fraction is loaded onto the iminobiotin affinity resin where the streptavidin is bound. The column is washed with equilibration buffer and the streptavidin is eluted with an ammonium acetate buffer, most preferably 50mM ammonium acetate at about pH 4.0. Streptavidin fractions are identified by U.V. detection and are further desalted and purified by standard gel filtration chromatography. Several gel filtration resins may be used but Sephacryl 200 is preferred. Final determination of the presence of pure streptavidin may be made by several methods, including for example, Western blot analysis.

The examples illustrate the isolation of the streptavidin gene, and the engineering of suitable closing vectors containing ligated DNA encoding operable expression elements and streptavidin. The examples also

describe the transformation of *Bacillus* host cells capable of expressing and secreting the streptavidin protein and the means whereby said streptavidin may be purified.

EXAMPLE 1

Construction of a *npr-sav* Hydrid Gene Fusion in *B. subtilis*

The *sav* gene was isolated as a 2 kb BamHI fragment by standard methods from *Streptomyces avidinii* and cloned in plasmid pSK (Stratagene, 11099 North Torrey Pines Road, La Jolla, CA 92037). An EcoRV site was engineered at the start site of the mature *sav* gene by site-directed mutagenesis and the resulting plasmid containing the *sav* gene was designated pBE651. The restriction map of plasmid pBE651 and the position of EcoRV site with reference to the mature coding region is illustrated in Figure 1a. Plasmid pBE83 (24) is an *E. coli-B. subtilis* phagemid vector containing the *npr* expression element and signal peptide fused to mature *Lvs*. *B. subtilis* strains containing pBE83 secrete Lvs. Colonies can be easily visualized on agar plates due to the formation of levan which is produced in the presence of sucrose contained in the agar. Plasmid pBE83 was digested with EcoRV and BamHI and ligated to the EcoRV-BamHI digested pBE651. *B. subtilis* strain BE1510 was transformed with the ligated DNA and plated on LB agar + 5% sucrose + chloramphenicol (5 ug/ml). A total of approximately 800 transformants were obtained and 128 *B. subtilis* clones that did not produce levansucrase were identified and were screened by colony immunoassay using commercially purchased anti-streptavidin antiserum. Four independent positive clones were obtained and designated as pBE659A, pBE659B, pBE659C and pBE659D. All further characterizations were carried out using plasmid pBE659A, which will be referred to as

pBE659. This plasmid has been deposited in the permanent culture collection of the American Type Culture Collection, Rockville, MD, 20852 and has been designated accession number ATCC 68977. The deposit was made in accord with the Budapest Treaty Requirements for purposes of patent procedure with access provisions as set forth under 37 C.F.R. 1.14 and 35 U.S.C. 122.

Construction *apr-sav*, *npr-sav*, *bar-sav* and *lvs-sav* Gene Fusions in *E. coli*

The following series of *sav* gene fusions were also constructed in *E. coli* and were subsequently transferred into *B. subtilis*. Plasmids pBE30, pBE90, pBE91 and pBE597 all contain the mature sequence of *phoA* fused to *apr*, *npr*, *bar* and *lvs* expression elements, respectively. These are shuttle phagemids containing origins of replication, of pBR322, pUB110 and M13K07 (24).

E. coli strains containing pBE30, pBE90, pBE91 and pBE597 secrete PhoA in *E. coli*. *PhoA* reacts with 5 bromo-4-chloro-3-indolyl phosphate (Sigma, St. Louis, MO) to provide a blue color on the indicator plates (23). Plasmids pBE30, pBE90, pBE91, pBE597 and pBE651 were digested with EcoRV and Pst 1 and separated on 1-2% low melting agarose. The large fragment from pBE30, pBE90, pBE91, pBE597 and the small fragment from pBE651 were cut and purified using Geneclean (P.O. Box 2284, La Jolla, CA 92038). The small *sav* fragment was ligated to each of the large fragments and *E. coli* strain XL1 (24) was transformed. The white transformants were screened by colony immunoassay using anti-sav antiserum (18). The plasmids were isolated, verified by methods well known to one skilled in the art and designated as pBE660(Apr-Sav), pBE661(Npr-Sav), pBE662(Bar-Sav) and pBE663(Lvs-Sav). *B. subtilis* strain BE1500 was transformed with pBE660, pBE661, pBE662 and pBE663 and Kan^r transformants were obtained. Colony immunoassay

revealed that all transformants secreted Sav. (Western blot analysis revealed that strains containing pBE660, pBE661, pBE662, and pBE663 secreted streptavidin into the growth medium.)

EXAMPLE 2

Expression of Sav in *B. subtilis*

B. subtilis strain containing pBE659A(*npr-sav*) and pBE20 (vector) (24), was grown in Medium B + chloramphenicol (5 ug/ml) and samples of the medium (1 ml) were withdrawn at 4, 6, 8, 10, 12 and 24 hours after the start of the culture (20). Samples were centrifuged in the presence of the 2mM phenyl methylsulfonyl fluoride (PMSF) for 5 min and the supernatant was respun and processed for Western analysis. Samples from pBE659(*npr-sav*) vector and commercially purchased streptavidin were separated on a 10 to 20% SDS-PAGE gel (Daiichi min gel, Integrated Separation Systems, MA 01136) followed by protein electroblotting onto a nitrocellulose filter and analyzed using anti-streptavidin antibody (Sigma, St. Louis, MO). Western blot analysis demonstrated that a protein band corresponding to mature-Sav (*npr-sav*) was present in pBE659(*npr-sav*) at 4 hours which increased in intensity with time and was present even after 24 hours (Figure 3). This band was absent in the strain that contained only the pBE20 vector. Thus, it is demonstrated that *B. subtilis* can secrete soluble Sav into the growth medium. It was noted that in this instance extracellular streptavidin accumulated to approximately 20 to 30 mg/liter after about 12 hours.

EXAMPLE 3

Purification of Streptavidin from the *B. subtilis* Culture Supernatant

The method of purification of streptavidin was based on the published methods described in references

(25), (26), and (27). *B. subtilis* strain BE1510 containing plasmid pBE659 was grown in 250 ml of the following medium: 0.6% Casaminoacids in 1X Castenholz medium + 1% glycerol + 0.01% yeast extract + 25 mM potassium phosphate buffer pH 7.0, 50 ug per ml of tryptophan, methionine and lysine and Cm (5 ug/ml) for 8 hours at 37°C [10X Castenholz basal stock contains per liter of distilled water nitriloacetic acid 1g; $CaSo_4 \cdot 2H_2O$, 0.6 g; $MgSO_4 \cdot 7H_2O$, 1.0; NaCl, 0.8 g; KNO_3, 1.03 g; $NaNO_3$, 6.89 g; Na_2HPO_4, 1.11 g; $FeCl_3$ solution 0.28 g/liter, 10 ml; Nitsch's Trace elements 10 ml; (Nitsch's Trace element Solution per liter of distilled water H_2SO_4, 0.5 ml; $MnSo_4 \cdot H_2O$, 2.2 g; $ZnSo_4 7H_2O$, 0.5 g; H_3BO_3, 0.5 g; $CuSO_4$, 0.016 g; $Na_2MoO_4 \cdot 2H_2O$, 0.025 g and $CoCl_2 \cdot 6H_2O$, 0.046 g)].

The pH of the medium was checked and maintained around pH 7.0 by the addition of sodium hydroxide. The bacteria were harvested after 8 hours and the growth medium was separated from the bacteria by centrifugation at 6,000 g for 20 min at 4°C in the presence of protease inhibitor (2 mM phenyl methyl sulfonyl fluoride). Ammonium sulfate was added to 70% to the supernatant and left stirring at 4°C overnight. The ammonium sulfate precipitate containing the partially purified streptavidin was collected by centrifugation at 6,000 g for 30 min and dissolved in 7 ml of 0.05M sodium bicarbonate buffer pH 11 + 0.5M NaCl and loaded onto a iminobiotin agarose column (5 ml). Iminobiotin agarose (Sigma) was prepared prior to sample loading by washing with 0.05M sodium bicarbonate buffer pH 11 + 0.5 M NaCl. After sample loading, the column was washed with 10 ml of 0.05 M sodium bicarbonate buffer pH 11 + 0.5 M NaCl and fractions were collected. The column was then eluted with 4 ml of 0.05 M ammonium acetate elution buffer pH 4.0 and fractions were collected. Absorbance

at 280 nm was measured for the various fractions as shown in Figure 4. The peak fraction after the pH shift on the iminobiotin column was subjected to gel filtration on a Sephacryl 200 column. The elution profile was consistent with an apparent molecular weight of approximately 60 kd suggesting that the tetrameric form of streptavidin had been purified. Samples from a few fractions as indicated in Figure 4 were analyzed by Western blot analysis. The results obtained in the Western blot were consistent with the properties of a tetrameric Sav, i.e., elution after shift in pH from iminobiotin column and a molecular weight of approximately 60 kd.

B. subtilis strain BE1510 containing pBE659 was grown in synthetic S7 medium and Sav was isolated using the above method or a batch method in which the iminobiotin agarose was added to the ammonium sulfate fraction followed by loading onto a column. Thus, Sav could be isolated from different media using methods that could be scaled up for downstream processing.

EXAMPLE 4

Role of N-terminal Residues 1 to 14 in the Secretion of Streptavidin

Construction of a gene fusion consisting of the *npr* signal peptide coding region fused to the 15th codon of mature streptavidin is described. The following steps were involved in the gene fusion construction: a) engineering the *B. amyloliquefaciens npr* secretion vector; b) engineering the *sav* mature sequence and c) fusion of the modified *sav* sequence to the modified *npr* vector.

(a) Engineering the *npr* Secretion Vector

The promoter and signal peptide coding region of *npr* gene was amplified from plasmid pBE80 (24) using primers

VN16 5'-ATGCATGGTACCGATCTAACATTTTCCCC-3' (SEQ ID NO:1) and VN47 5'-GACGTATAT**GATATC**CGC**GCTAGC**ACCCGGCAGACTGAT-3' (SEQ ID NO:2) using the PCR method according to the manufacturer's instructions (Perkin Elmer Cetus , 761 Main Ave., Norwalk, CT 06859, Gene Amp PCR kit). The amplified fragment was treated with Klenow fragment to fill in any ragged ends and digested with Kpn and EcoRV, and was then purified using Geneclean kit (P.O. Box 2284, La Jolla, CA 92038). Plasmid pBE592 was the source of the vector containing *E. coli* alkaline phosphatase and is similar to plasmid pBE597 (24) except that it contains a seven amino acid deletion in the signal peptide. *E. coli* alkaline phosphatase activity is indicative of export of the protein and colonies that secrete phosphatase appear blue on indicator plates (LB agar + 5 bromo-4-chloro-3-indolyl phosphate) (23). *E. coli* strain containing plasmid pBE592 is white on LB agar + 5 bromo-4-chloro-3-indolyl phosphate because phosphatase is not secreted. The PCR amplified pBE80 fragment (Kpn-EcoRV) was ligated to pBE592 digested with Kpn-EcoRV and *E. coli* was transformed with the ligated DNA and plated on a LB plate + 100 ug/ml ampicillin + 50 ug/ml 5 bromo-4-chloro-3-indolyl phosphate (Sigma). Blue colonies were isolated, verified by restriction analysis, and one such plasmid was designated by pBE93.

(b) Engineering of sav Gene

Single stranded DNA was isolated from pBE651 and site directed mutagenesis was performed using two oligonucleotides VN44 and VN51. Oligonucleotide VN44 (GTC TCG GCC GCC GAG **GCT AGC GCC** GCC GGC ATC ACC GGD) (SEQ ID NO: 3) codes for a Nhe 1 site which precedes codon 15 of Sav. VN51 (GAC ACC TTC ACC AAG GTG **TAG GTC GAC** AAG CCG TCC GCC) (SEQ ID NO:4) codes for a translational terminator and a Sal 1 site downstream of

codon 133 of mature sav. The transformants were screened primarily for the presence of Nhe 1 site followed by screening for the Sal 1 site. pBE670 contained the newly engineered Nhe 1 site; however, Sal I digestion resulted in a partial digestion suggesting that it had both the parent (without a Sal site) and the mutant plasmids in the cell.

(c) Fusion of the Modified *sav* Sequence to the Modified *npr* Vector

Plasmid pBE93 was digested with Nhe and Pst and ligated to Nhe 1-Pst digested pBE670, and *E. coli* was then transformed and screened for streptavidin production by colony immunoassay. One of the positive clones was designated as plasmid pBE673. Restriction analysis of pBE673 revealed that it did not contain the Sal site at the 3' end and thus it encoded for a Sav protein consisting of residues 15 to 159. *B. subtilis* BE1500 was transformed with pBE673 and Kan R transformants were screened by colony immunoassay for streptavidin production. *B. subtilis* strain BE1500(pBE673) and BE1510(pBE659) were grown in medium A + kanamycin (10 ug/ml) or chloramphenicol (5 ug/ml). The extracellular streptavidin was analyzed by Western blot analysis (Figure 5b). The mobility of streptavidin produced by pBE673 was faster than that of pBE659 due to the deletion of residues 1 to 14 of mature streptavidin. *B. subtilis* strain containing pBE673 was able to secrete streptavidin efficiently into the growth medium suggesting that the hybrid fusion junction of the *npr* signal peptide fused to the 15th residue of streptavidin was efficiently recognized by *B. subtilis*.

The DNA sequence across the fusion junction in pBE673 is shown below.

npr-sav(15-159)

Sequence at the N-terminus of sav

Signal(*npr*) Nhe I +15(*sav*)

ATC AGT CTG CCG GGT GCT AGC GCC *GCC GGC ATC*

(SEQ ID NO:5)

EXAMPLE 5

Vectors to Make C-terminal Fusions

This example describes the construction of a set of vectors that can be used to make a variety of C-terminal fusions. The rationale for the construction of these vectors is at least two-fold. Firstly, the role of the C-terminal tail in the oligomerization of streptavidin in the growth medium can be studied. Secondly, one can identify the most stable bifunctional molecule that can maintain both the biotin binding feature of the Sav and contain an additional enzymatic activity such as levansucrase, alkaline phosphatase, β-lactamase, protein A and luciferase. Fusion proteins are normally rapidly clipped at the fusion junction in the *B. subtilis* growth medium and thus by creating different fusion junctions, the most stable protein can be identified.

The restriction enzyme site that was engineered was Msc 1 because it generates a blunt end and is unique in the vectors. Msc 1 site was created by a six base (TGG CCA) insertion between mature-sav codons 132 and 133, 138 and 139, 145 and 146 and 159 and terminator codon by site-directed mutagenesis. The newly created Msc 1 site can be used to create translation fusions to any heterologous protein.

VN66 to create Msc 1 site at between codons 133 and 134.

ACC TTC ACC AAG GTG TGG CCA AAG CCG TCC GCC

(SEQ ID NO:6)

VN67 to create Msc 1 site between codons 138 and 139.
CCG TCC GCC GCC **TGG CCA** TCC ATC GAC GCG GCG
(SEQ ID NO:7)

VN68 to create Msc 1 site between codons 145 and 146.
GAC GCG GCG AAA AAG **TGG CCA** GCC GGC GTC AAC AAC
(SEQ ID NO:8)

VN62 to create Msc 1 site at the end of sav(159).
GAC GCC GTT CAG CAG **TGG CCA** TAG TCG CGT CCC GGC
(SEQ ID NO:9)

Plasmids with the Msc 1 site were identified by restriction analysis and designated as pBE626 (VN62), pBE627 (VN66), pBE628 (VN67), and pBE629 (VN68), respectively. *E. coli* strains containing pBE626, pBE627, pBE628, and pBE629 were positive for streptavidin production as determined by colony immunoassay. These vectors can be used to fuse any reporter protein (levansucrase, β-lactamase, PhoA, protein A or luciferase). Thus, the most stable protein that has both the biotin binding activity and the reporter protein can be obtained. Schematic representation of the relevant hybrid fusions are shown in Figure 6.

EXAMPLE 6

Secretion of Streptavidin as a Fusion Protein

This example describes the construction of a hybrid fusion protein containing streptavidin and levansucrase using a different signal peptide. Single stranded DNA from plasmid pBE651 (Figure 1a) was used to create an EcoRV site between codons 159 and the translational terminator of *sav* resulting in plasmid pBE653. Thus the *sav* gene can be isolated as an EcoRV fragment from

plasmid pBE653. Plasmid pBE311 codes for levansucrase and contains an EcoRV site between the second and third codon of mature levansucrase (24). Plasmid pBE311 was digested with EcoRV and ligated to the EcoRV digested pBE653. In order to identify clones coding for the fusion protein, *B. subtilis* transformants were screened by colony immunoassay using anti-streptavidin antiserum. pBE653 contains *sav* gene but *sav* is not expressed. pBE311 does not contain *sav* gene and therefore only the recombinants will express the fusion protein. Plasmids containing the *sav* fragment fused to levanscurase were designated as pBE655. *B. subtilis* strains containing pBE655 were grown and labelled with ^{35}S-methionine and the medium was analyzed by gel electrophoresis for the presence of streptavidin and levansucrase. Upon analysis, proteins corresponding to the fusion protein (Sav-Lvs) and mature levansucrase (Lvs) were found to be present in the growth medium. *B. subtilis* strains pBE655 and pBE311 were grown in medium A (18), and the extracellular levansucrase activity was measured (20). Levansucrase activity in the two independent transformants from pBE655A and pBE655B were found to be 39% and 86% higher, respectively, than those of pBE311.

REFERENCES

1. Weber, P. C., Ohlendorf, D. H., Wendolski, J. J., and Salemme, F. R., (1989) Science 243:85-88.

2. Argarana, C. E., Kuntz, I. D., Birken, S., Axel R., and Cantor, C. R., (1986) Nucleic Acids Research 14:1871-1882.

3. Meade, H., Garwin, J. PCT Patent Application WO86/02077.

4. Cantor C., Axel, R., and Argarana, C., (1989) U.S. Patent 4,839,293. Claims 1-25 withdrawn due to Inteference Decision No. 102,477.

5. Sano, T., and Cantor, C. R., (1990) Proc. Natl. Acad. Sci. 87:142-146.

6. Sano, T., and Cantor, C. R., (1991) Biochemical and Biophysical Research Communications 176(2):571-577.

7. Chang, S., (1987) U.S. Patent 4,711,843.

8. Nagarajan, V., Rhodes, C. S., and Banner, C. D. B., (1989) U.S. Patent 4,801,537.

9. Kovacevic, S., Miller, J. R., and Hsuiung, H. M., (1988) U.S. Patent 4,783,405.

10. Stephens et al., U.S. Patent 4,769,327.

11. Biotechnology Handbook 2, Bacillus, C. R. Hardwood, Ecf., Plenum Press, New York (1989).

12. Palva et al., (1982), Proc. Natl. Acad. Sci., USA 79, 5582-5586.

13. Palva et al., (1983), Gene 22, 229-235.

14. Dion et al., (1989), Biochime, 71, 747-7555.

15. Nagarajan, V., et al., U.S. Patent Application Serial No. 07/803,843 and PCT Application US91/05909.

16. Sambrook, J., Fritsch, E. F., Maniatis, T., (1989) Molecular Cloning Volumes 1, 2 and 3, Cold Spring Harbor Laboratory Press, New York.

17. Goeddel, D. V., (1990) Methods in Enzymology 185; Gene Expression Technology Academic Press, New York.

18. Nagarajan, V., (1990) Methods in Enzymology 185 214-223.

19. Vasantha, N., Thompson, L. D., Rhodes, C., Banner, C., Nagle, J., and Filpula, D., (1984) J. Bacteriol. 159 811-189.

20. Tang, Lr., Lenstra, R., Borchert, T. V., and Nagarajan, V., (1990) Gene 96:89-93.

21. Chang, C. N., Kuang, W. J., Chen, E. Y., (1986) Gene 44, 121-126.

22. Paddon, C. J. and Hartley, R. W., (1985) Gene 40:231-239.

23. Harwood, C. R., and Cutting, S. N., (1990) Molecular Biological Methods for Bacillus, John Wiley and Sons, New York.

24. Nagarajan, V., Albertson, H., Chen, N., and Ribbe, J., (1992) Modular Expression and Secretion Vectors for *Bacillus subtilis*, Gene, in press.

25. Hofmann, K., Wood, S. W., Brinton, C. C., Montibeller, J. A., and Finn, F. M., (1980) Proc. Natl. Acad. Sci. 77(8)4666-4668.

26. Heney, G., and Orr, G. A., (1981) Analytical Biochemistry 114:92-96.

27. Cuatrecasas, P., and Wilchek, M., (1968), Biochemical and Biophysical Research Communications 33(2)235-239.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: NAGARAJAN, VASANTHA

(ii) TITLE OF INVENTION: PRODUCTION OF STREPTAVIDIN FROM BACILLUS SUBTILLIS

(iii) NUMBER OF SEQUENCES: 14

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: DU PONT COMPANY

(B) STREET: BARLEY MILL PLAZA 36

(C) CITY: WILMINGTON

(D) STATE: DELAWARE

(E) COUNTRY: USA

(F) ZIP: 19880-0036

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE:

(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: GEIGER, KATHLEEN W

(C) REFERENCE/DOCKET NUMBER: CR 9029

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 302-892-2118

(B) TELEFAX: 302-892-7949

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGCATGGTA CCGATCTAAC ATTTTCCCC 29

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GACGTATATG ATATCCGCGC TAGCACCCGG CAGACTGAT 39

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTCTCGGCCG CCGAGGCTAG CGCCGCCGGC ATCACCGGC 39

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GACACCTTCA CCAAGGTGTA GGTCGACAAG CCGTCCGCC 39

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATCAGTCTGC CGGGTGCTAG CGCCGCCGGC ATC 33

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ACCTTCACCA AGGTGTGGCC AAAGCCGTCC GCC 33

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCGTCCGCCG CCTGGCCATC CATCGACGCG GCG 33

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GACGCGGCGA AAAAGTGGCC AGCCGGCGTC AACAAC 36

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GACGCCGTTC AGCAGTGGCC ATAGTCGCGT CCCGGC 36

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GTTCAGGCCG CTGAGGATAT C 21

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCTTCGGCAG ATATCTCC 18

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTTCAGGCCG CTGAGGATAT CTCC 24

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 552 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ATGCGCAAGA TCGTCGTTGC AGCCATCGCC GTTTCCCTGA CCACGGTCTC GATTACGGCC 60

AGCGCTTCGG CAGACCCCTC CAAGGACTCG AAGGCCCAGG TCTCGGCCGC CGAGGCCGGC 120

ATCACCGGCA CCTGGTACAA CCAGCTCGGC TCGACCTTCA TCGTGACCGC GGGCGCCGAC 180

GGCGCCCTGA CCGGAACCTA CGAGTCGGCC GTCGGCAACG CCGAGAGCCG CTACGTCCTG 240

ACCGGTCGTT ACGACAGCGC CCCGGCCACC GACGGCAGCG GCACCGCCCT CGGTTGGACG 300

GTGGCCTGGA AGAATAACTA CCGCAACGCC CACTCCGCGA CCACGTGGAG CGGCCAGTAC 360

GTCGGCGGCG CCGAGGCGAG GATCAACACC CAGTGGCTGC TGACCTCCGG CACCACCGAG 420

GCCAACGCCT GGAAGTCCAC GCTGGTCGGC CACGACACCT TCACCAAGGT GAAGCCGTCC 480

GCCGCCTCCA TCGACGCGGC GAAGAAGGCC GGCGTCAACA ACGGCAACCC GCTCGACGCC 540

GTTCAGCAGT AG 552

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 183 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Arg Lys Ile Val Val Ala Ala Ile Ala Val Ser Leu Thr Thr Val
1 5 10 15

Ser Ile Thr Ala Ser Ala Ser Ala Asp Pro Ser Lys Asp Ser Lys Ala
20 25 30

Gln Val Ser Ala Ala Glu Ala Gly Ile Thr Gly Thr Trp Tyr Asn Gln
35 40 45

Leu Gly Ser Thr Phe Il Val Thr Ala Gly Ala Asp Gly Ala Leu Thr
50 55 60

Gly Thr Tyr Glu Ser Ala Val Gly Asn Ala Glu Ser Arg Tyr Val Leu
65 70 75 80

Thr Gly Arg Tyr Asp Ser Ala Pro Ala Thr Asp Gly Ser Gly Thr Ala
85 90 95

Leu Gly Trp Thr Val Ala Trp Lys Asn Asn Tyr Arg Asn Ala His Ser
100 105 110

Ala Thr Thr Trp Ser Gly Gln Tyr Val Gly Gly Ala Glu Ala Arg Ile
115 120 125

Asn Thr Gln Trp Leu Leu Thr Ser Gly Thr Thr Glu Ala Asn Ala Trp
130 135 140

Lys Ser Thr Leu Val Gly His Asp Thr Phe Thr Lys Val Lys Pro Ser
145 150 155 160

Ala Ala Ser Ile Asp Ala Ala Lys Lys Ala Gly Val Asn Asn Gly Asn
165 170 175

Pro Leu Asp Ala Val Gln Gln
180

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on page 18, lines 1-4 of the description

A. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet ☐

Name of depositary institution

AMERICAN TYPE CULTURE COLLECTION

Address of depositary institution (including postal code and country)

12301 Parklawn Drive
Rockville, Maryland 20852
US

Date of deposit	Accession Number
08 May 1992 (08.05.92)	68977

B. ADDITIONAL INDICATIONS (leave blank if not applicable). This information is continued on a separate attached sheet ☐

In respect of those designations in which a European patent is sought, a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample. (Rule 28(4) EPC)

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

D. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. ☐ This sheet was received with the international application when filed (to be checked by the receiving Office)

(Authorized Officer)

☐ The date of receipt (from the applicant) by the International Bureau

was

(Authorized Officer)

Form PCT/RO/134 (January 1981)

What is claimed is:

1. A method for producing tetrameric, biologically active streptavidin by secretion from *Bacillus subtilis*, comprising:

a. transforming *Bacillus subtilis* with a gene construct comprising a sequence encoding streptavidin operably linked to a sequence encoding a signal peptide and an expression element, wherein said sequence encoding a signal peptide is isolated from DNA encoding an exoprotein of bacteria and said expression element is isolated from DNA encoding a gram positive bacterial protein;

b. growing the transformed *Bacillus subtilis* in suitable growth medium whereby streptavidin is secreted into the growth medium; and

c. purifying the streptavidin from the growth medium.

2. The method of Claim 1 wherein said sequence encoding a signal peptide and said expression element are each independently selected from the group consisting of *apr*, *npr*, *lvs* and *bar* derived from *Bacillus amyloliquefaciens*.

3. The method of Claim 1 wherein said sequence encoding a signal peptide and said expression element are derived from *npr* from *Bacillus amyloliquefaciens*.

4. The method of Claim 1 wherein said sequence encoding a signal peptide and said expression element are derived from *lvs* from *Bacillus amyloliquefaciens*.

5. The method of Claim 1 wherein said expression element is comprised of a promoter sequence and a ribosomal binding sit sequence.

6. The method of Claim 5 wherein said promoter sequence is inducible.

7. A transformed *Bacillus subtilis* bacterium capable of secreting tetrameric, biologically active streptavidin having a gene construct comprising a sequence encoding streptavidin operably linked to a sequence encoding a signal peptide and an expression element, wherein said sequence encoding a signal peptide is isolated from DNA encoding an exoprotein of bacteria and said expression element is isolated from DNA encoding a gram positive bacterial protein.

8. A bacterium of Claim 7 wherein said sequence encoding a signal peptide and said expression element are each independently selected from the group consisting of *apr*, *npr*, *lvs* and *bar* genes derived from *Bacillus amyloliquefaciens*.

9. A method for producing a fused gene product comprised of tetrameric, biologically active streptavidin fused to a second desired protein, by secretion from *Bacillus subtilis*, comprising:

a. transforming *Bacillus subtilis* with a fused gene construct comprising a sequence encoding streptavidin fused to a sequence encoding a second desired protein, wherein said fused sequence is operably linked to a sequence encoding a signal peptide and an expression element, and wherein said sequence encoding a signal peptide is isolated from DNA encoding an exoprotein of bacteria and said expression element is

isolated from DNA encoding a gram positive bacterial protein;

b. growing the transformed *Bacillus subtilis* in suitable growth medium whereby the fused streptavidin and desired protein are secreted into the growth medium; and

c. purifying the fused streptavidin and desired protein from the growth medium.

10. The method of Claim 9 wherein said sequence encoding a signal peptide and said expression element are each independently selected from the group consisting of *apr*, *npr*, *lvs* and *bar* genes from *Bacillus amyloliquefaciens*.

11. The method of Claim 9 wherein said fused gene product is comprised of tetrameric, biologically active streptavidin fused to the C-terminus of the second desired protein.

12. The method of Claim 9 wherein said fused gene product is comprised of tetrameric, biologically active streptavidin fused to the N-terminus of the second desired protein.

13. The method of Claim 9 wherein said second target protein is selected from the group consisting of levansucrase, alkaline phosphatase, β-lactamase, luciferase, and Protein A.

14. The method of Claim 12 wherein the fused gene product is comprised of streptavidin at the N-terminal region of the fused protein and levansucrase at the C-terminal region of the fused protein; and wherein said fused gene construct is comprised of the expression

element and sequence encoding the signal peptide isolated from the levansucrase gene of *Bacillus amyloliquefaciens*, fused to a sequence encoding mature streptavidin, fused to a sequence encoding mature levansucrase from *Bacillus amyloliquefaciens*.

15. A fused gene product produced by the method of Claim 11 or 12.

16. A transformed *Bacillus subtilis* bacterium capable of secreting tetrameric, biologically active streptavidin fused to a second desired protein; said bacterium having a fused gene construct comprising a sequence encoding streptavidin fused to a sequence encoding a second desired protein; wherein said fused sequence is operably linked to a sequence encoding a signal peptide and an expression element, and wherein said sequence encoding a signal peptide is isolated from DNA encoding an exoprotein of bacteria and said expression element is isolated from DNA encoding a gram positive bacterial protein.

17. The bacterium of Claim 15 wherein said fused protein is comprised of tetrameric, biologically active streptavidin fused to the C-terminus of the second desired protein.

18. The bacterium of Claim 15 wherein said fused protein is comprised of tetrameric, biologically active streptavidin fused to the N-terminus of the second desired protein.

19. The bacterium of Claim 16 wherein said sequence encoding a signal peptide and said expression element are each independently selected from the group

consisting of *apr*, *npr*, *lvs* and *bar* genes from *Bacillus amyloliquefaciens*.

20. The bacterium of Claim 16 wherein said second target protein is selected from the group consisting of levansucrase, alkaline phosphatase, β-lactamase, luciferase, and Protein A.

21. Plasmids pBE659, pBE660, pBE661, pBE662, pBE663, pBE673, and pBE655.

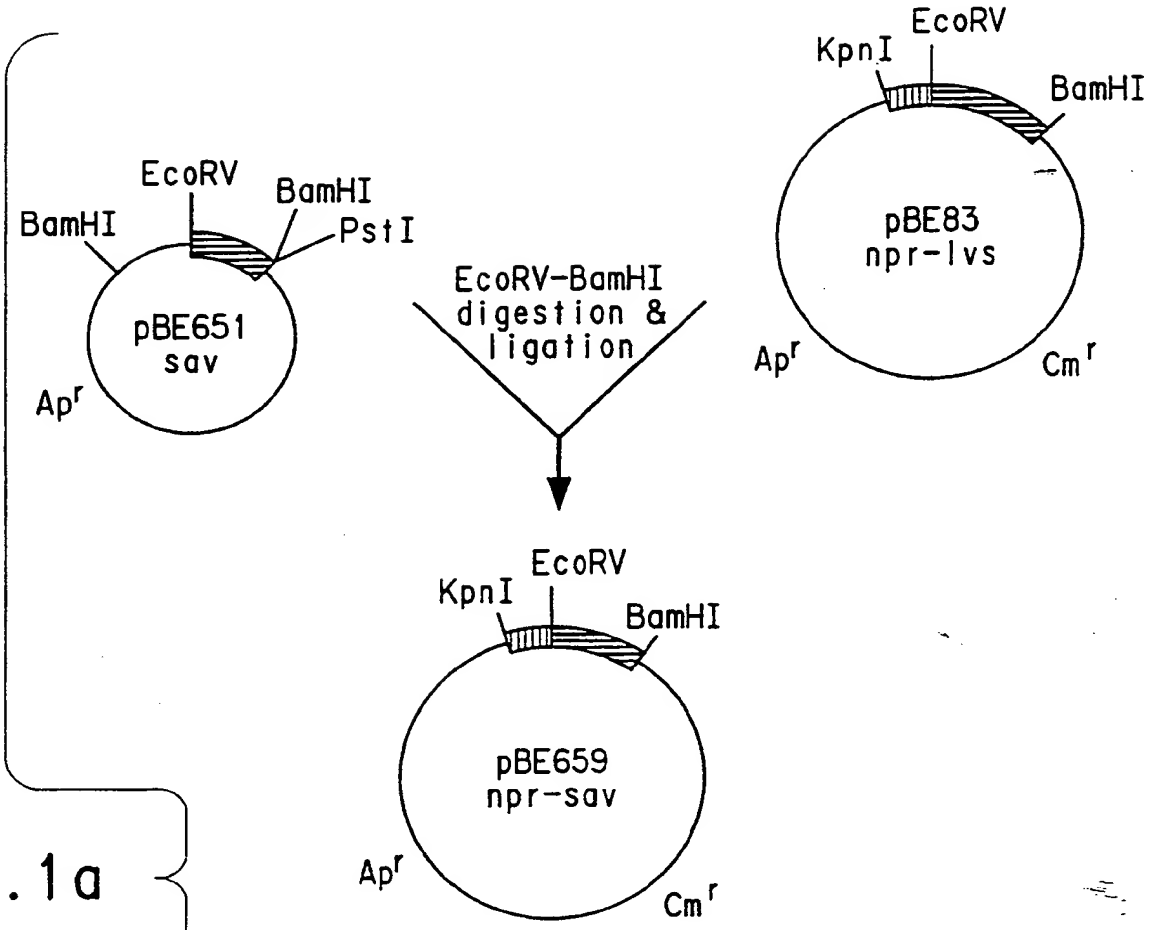


FIG.1a

A) DNA sequence across *npr* signal peptide cleavage site in pBE83.

```
-3      -1 | +1      EcoRV
GTT CAG GCC | GCT GAG GAT ATC
```

SEQ ID NO.: 10

B) DNA sequence across *sav* mature sequence in pBE651.

```
-3      -1 | +1  +2 +3
GCT TCG GCA | GAT ATC TCC
```

SEQ ID NO.: 11

C) DNA sequence across *npr-sav* fusion junction in pBE659.

```
-3      -1 | +1      EcoRV
GTT CAG GCC | GCT GAG GAT ATC TCC
```

SEQ ID NO.: 12

FIG. 1b

```
1
ATG CGC AAG ATC GTC GTT GCA GCC ATC GCC GTT TCC CTG ACC ACG GTC TCG ATT ACG GCC
Met arg lys ile val val ala ala ile ala val ser leu thr thr val ser ile thr ala
61     signal    GAT ATC                                              13
AGC GCT TCG GCA GAC CCC TCC AAG GAC TCG AAG GCC CAG GTC TCG GCC GCC GAG GCC GGC
ser ala ser ala asp pro ser lys asp ser lys ala gln val ser ala ala glu ala gly
121
ATC ACC GGC ACC TGG TAC AAC CAG CTC GGC TCG ACC TTC ATC GTG ACC GCG GGC GCC GAC
ile thr gly thr trp tyr asn gln leu gly ser thr phe ile val thr ala gly ala asp
181
GGC GCC CTG ACC GGA ACC TAC GAG TCG GCC GTC GGC AAC GCC GAG AGC CGC TAC GTC CTG
gly ala leu thr gly thr tyr glu ser ala val gly asn ala glu ser arg tyr val leu
241
ACC GGT CGT TAC GAC AGC GCC CCG GCC ACC GAC GGC AGC GGC ACC GCC CTC GGT TGG ACG
thr gly arg tyr asp ser ala pro ala thr asp gly ser gly thr ala leu gly trp thr
301
GTG GCC TGG AAG AAT AAC TAC CGC AAC GCC CAC TCC GCG ACC ACG TGG AGC GGC CAG TAC
val ala trp lys asn asn tyr arg asn ala his ser ala thr thr trp ser gly gln tyr
361
GTC GGC GGC GCC GAG GCG AGG ATC AAC ACC CAG TGG CTG CTG ACC TCC GGC ACC ACC GAG
val gly gly ala glu ala arg ile asn thr gln trp leu leu thr ser gly thr thr glu
421
GCC AAC GCC TGG AAG TCC ACG CTG GTC GGC CAC GAC ACC TTC ACC AAG GTG AAG CCG TCC
ala asn ala trp lys ser thr leu val gly his asp thr phe thr lys val lys pro ser
481 138
GCC GCC TCC ATC GAC GCG GCG AAG AAG GCC GGC GTC AAC AAC GGC AAC CCG CTC GAC GCC
ala ala ser ile asp ala ala lys lys ala gly val asn asn gly asn pro leu asp ala
541
GTT CAG CAG TAG      SEQ ID NO.:  13
val gln gln AMB      SEQ ID NO.:  14
```

FIG.2

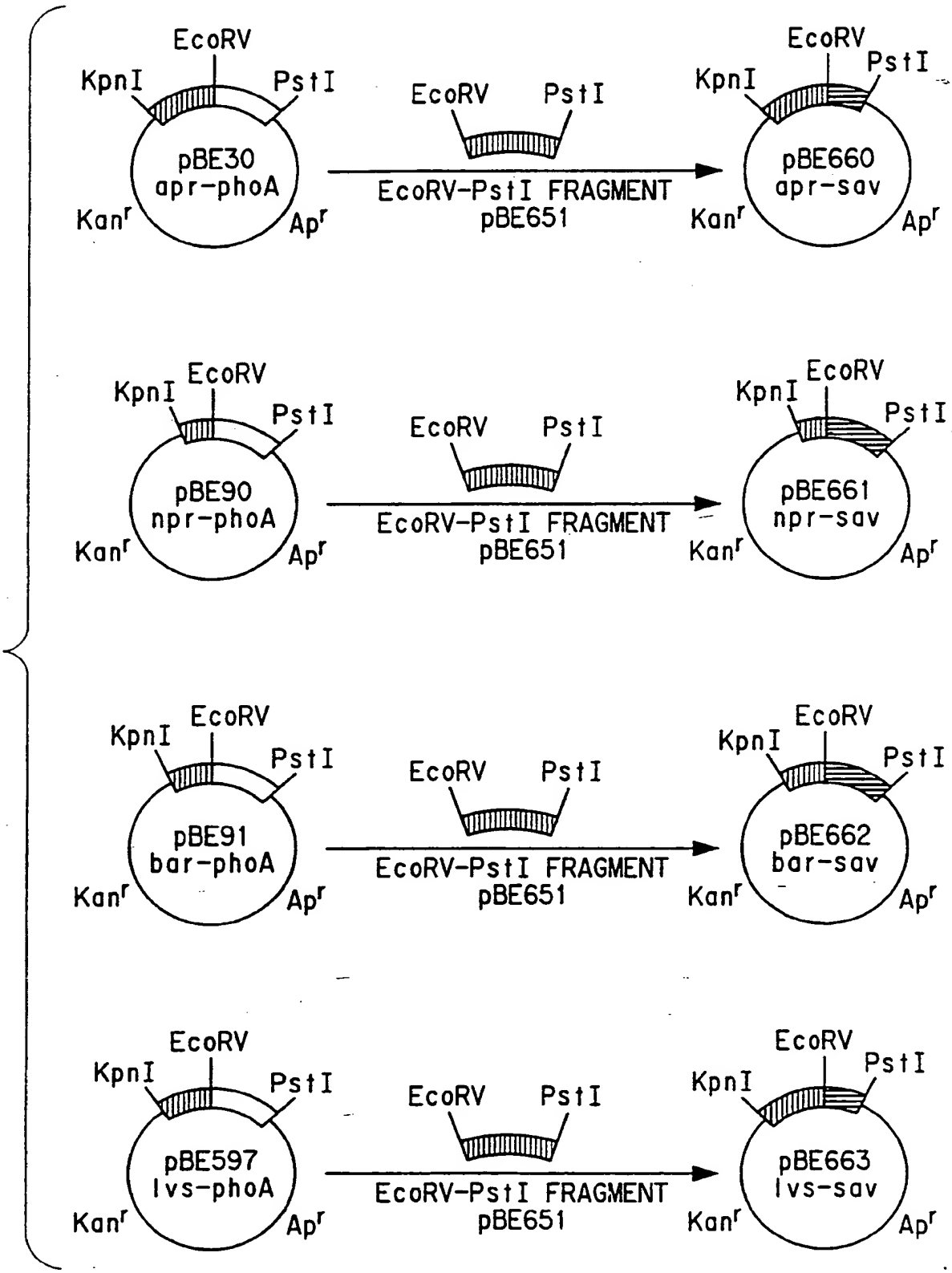

EcoRV
KpnI
PstI
pBE30
apr-phoA
Kan^r
Ap^r
EcoRV
PstI
EcoRV-PstI FRAGMENT
pBE651
EcoRV
KpnI
PstI
pBE660
apr-sav
Kan^r
Ap^r
EcoRV
KpnI
PstI
pBE90
npr-phoA
Kan^r
Ap^r
EcoRV
PstI
EcoRV-PstI FRAGMENT
pBE651
EcoRV
KpnI
PstI
pBE661
npr-sav
Kan^r
Ap^r
EcoRV
KpnI
PstI
pBE91
bar-phoA
Kan^r
Ap^r
EcoRV
PstI
EcoRV-PstI FRAGMENT
pBE651
EcoRV
KpnI
PstI
pBE662
bar-sav
Kan^r
Ap^r
EcoRV
KpnI
PstI
pBE597
lvs-phoA
Kan^r
Ap^r
EcoRV
PstI
EcoRV-PstI FRAGMENT
pBE651
EcoRV
KpnI
PstI
pBE663
lvs-sav
Kan^r
Ap^r

FIG.3

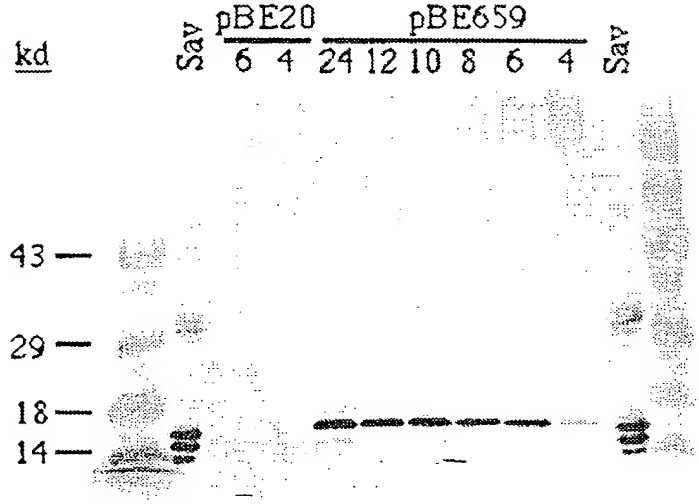


FIG.4a

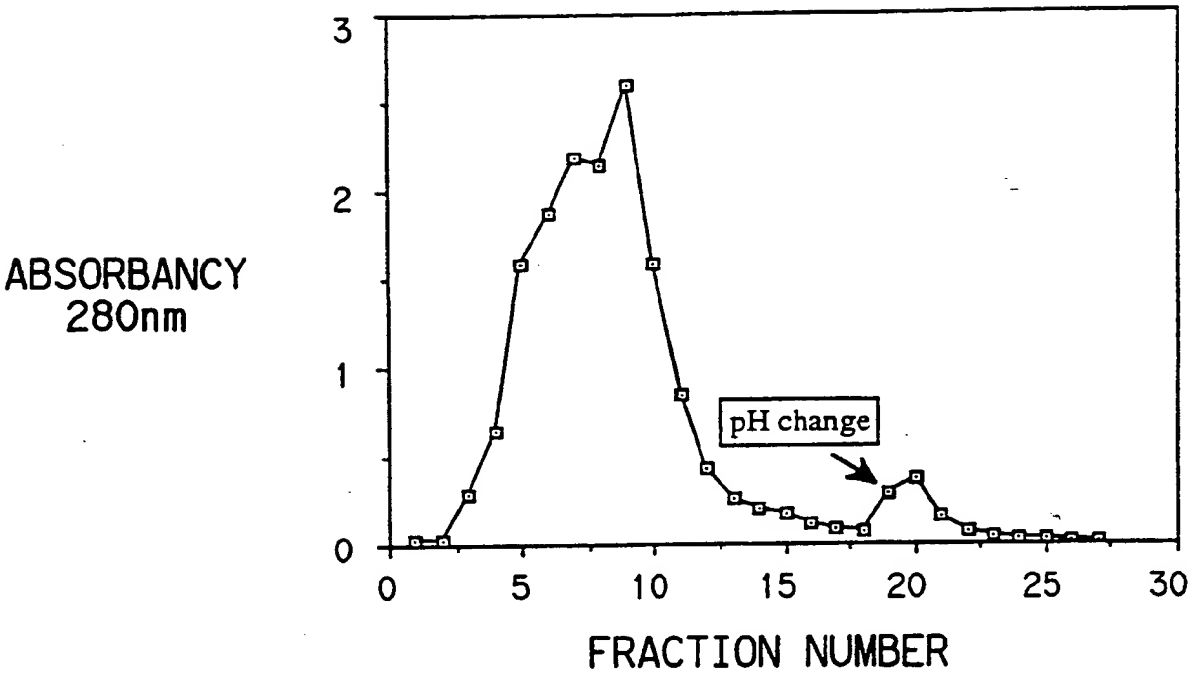

ABSORBANCY
280nm
3
2
1
0
0
5
10
15
20
25
30
pH change
FRACTION NUMBER

FIG.4B

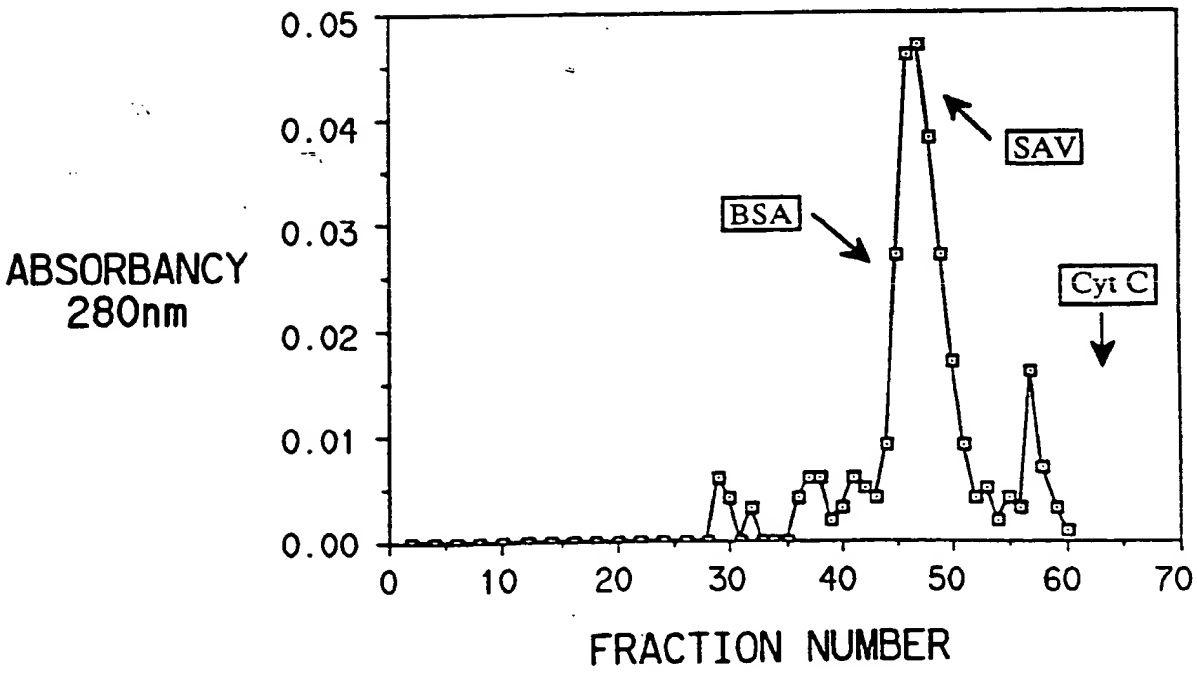

ABSORBANCY
280nm
0.05
0.04
0.03
0.02
0.01
0.00
0
10
20
30
40
50
60
70
SAV
BSA
Cyt C
FRACTION NUMBER

FIG. 4c

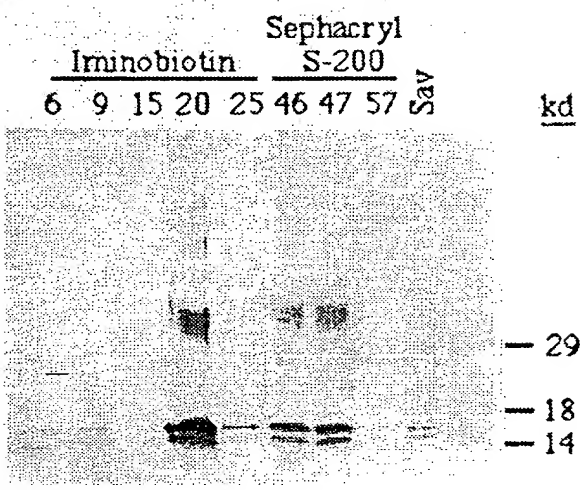


FIG.5a

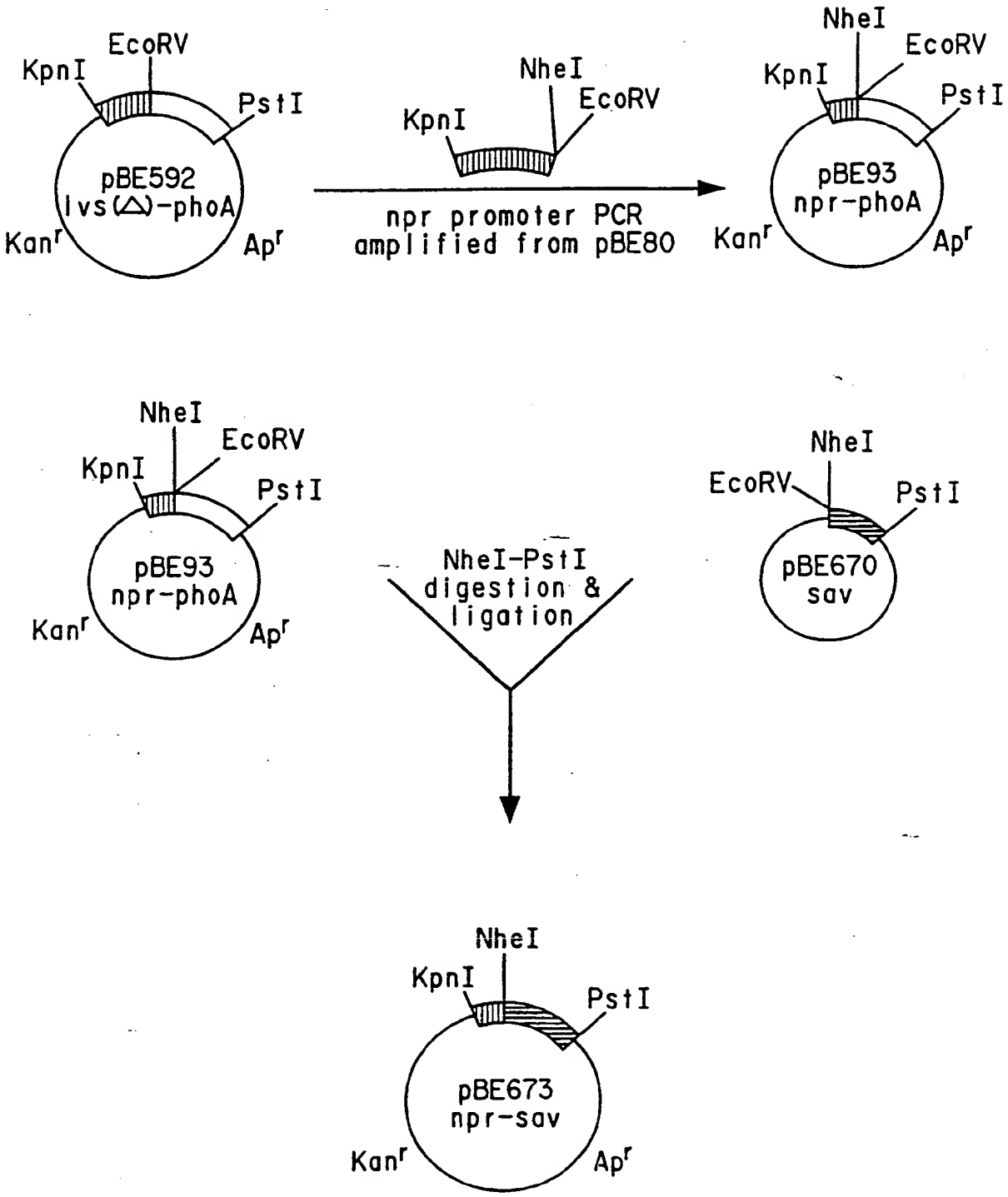

EcoRV
KpnI
PstI
pBE592
lvs(Δ)-phoA
Kan^r
Ap^r
NheI
EcoRV
KpnI
npr promoter PCR
amplified from pBE80
NheI
EcoRV
KpnI
PstI
pBE93
npr-phoA
Kan^r
Ap^r
NheI
EcoRV
KpnI
PstI
pBE93
npr-phoA
Kan^r
Ap^r
NheI-PstI
digestion &
ligation
NheI
EcoRV
PstI
pBE670
sav
NheI
KpnI
PstI
pBE673
npr-sav
Kan^r
Ap^r

FIG.5b

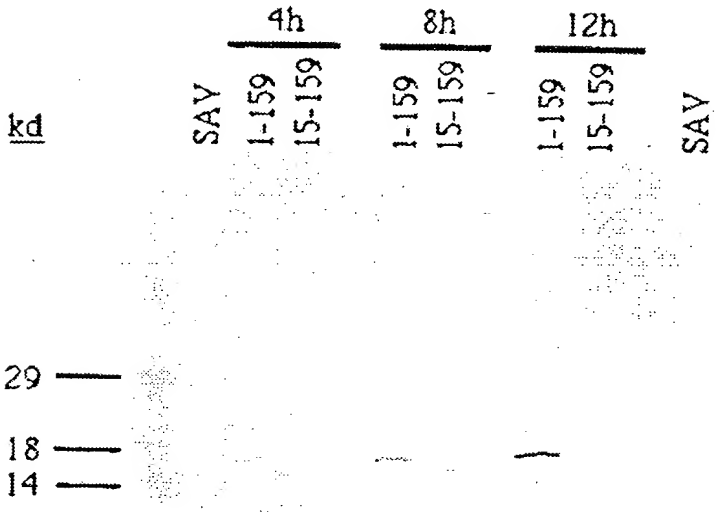


FIG.6

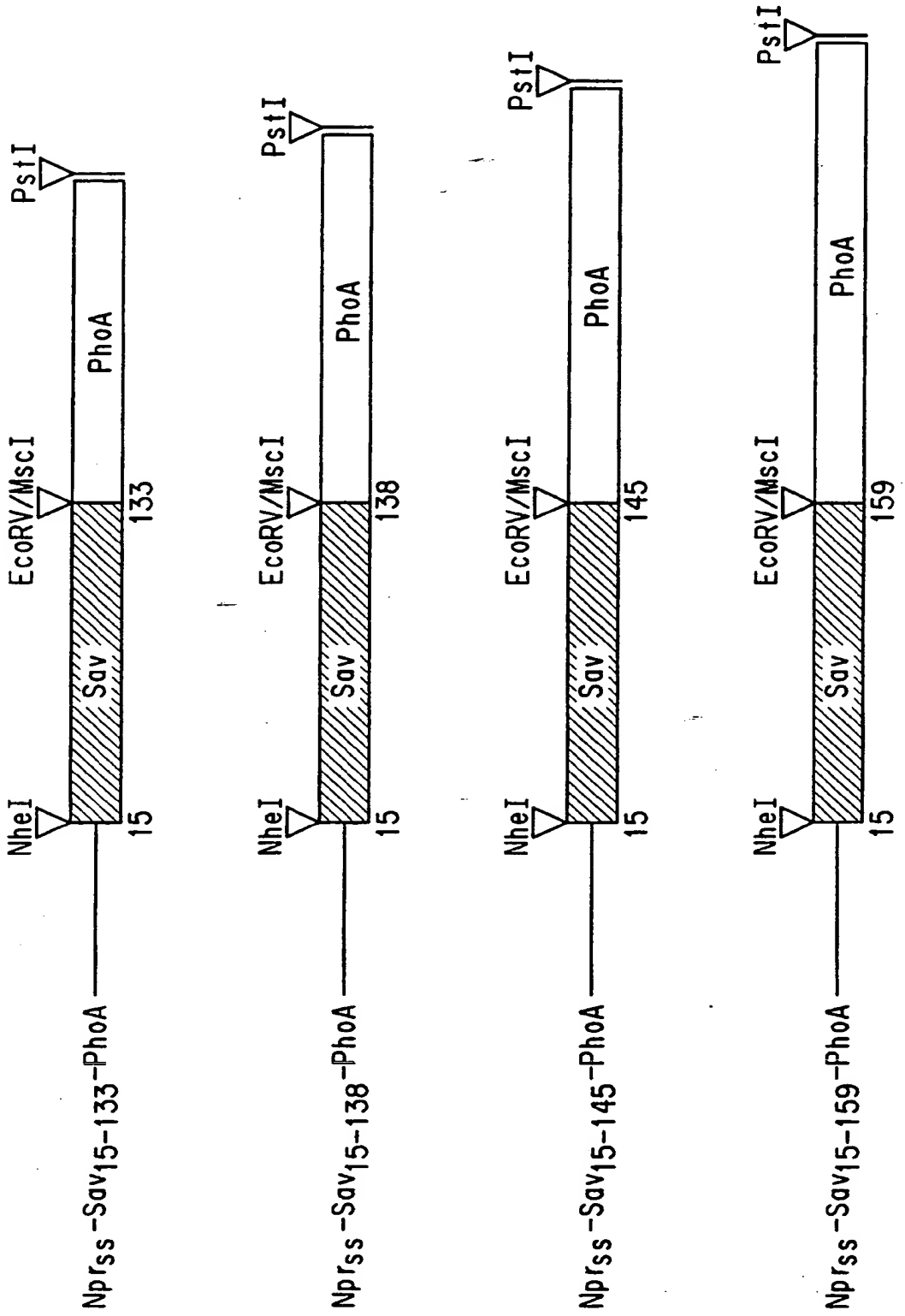
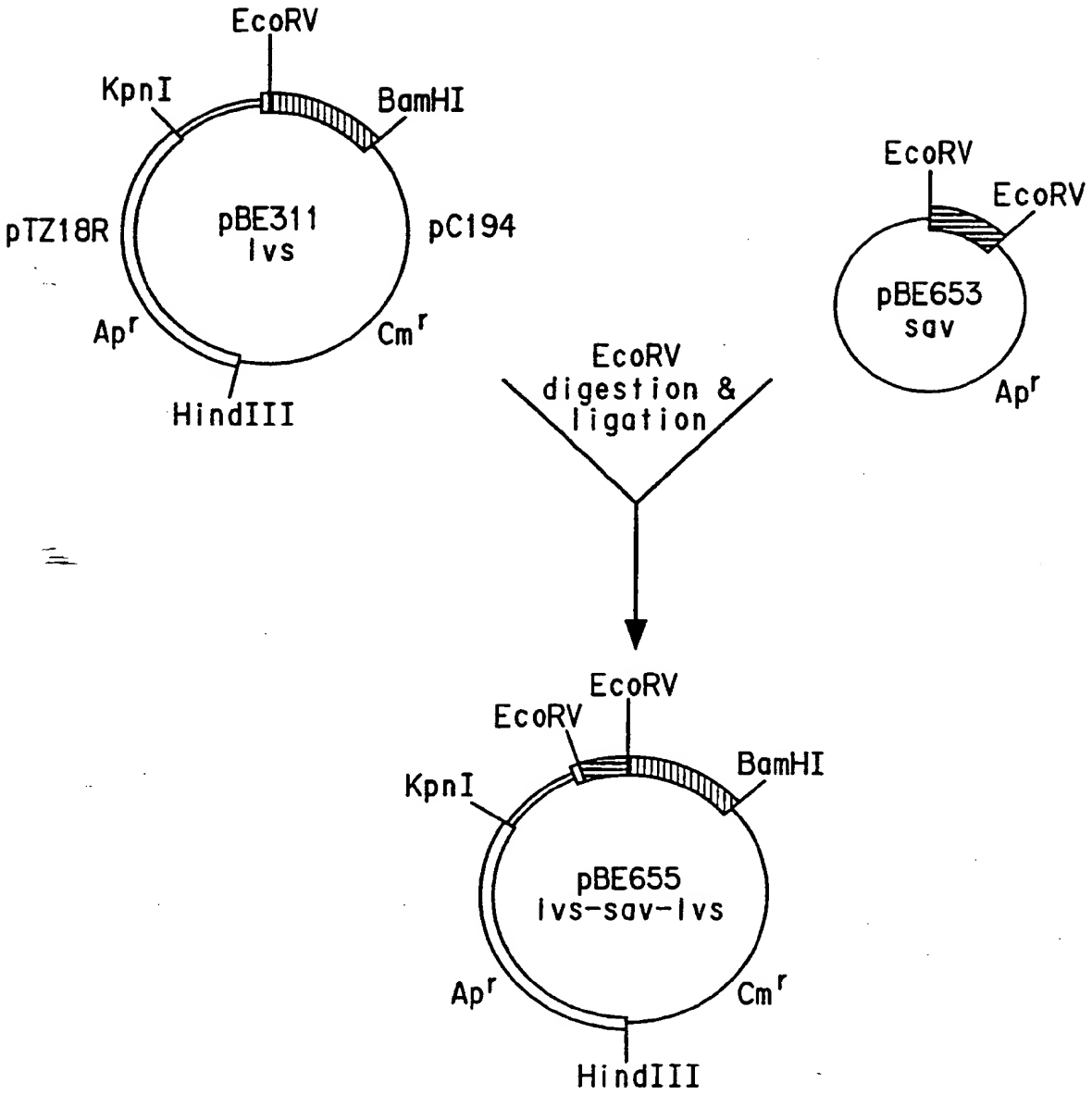

PstI
PhoA
EcoRV/MscI
133
Sav
NheI
15
Npr$_{ss}$-Sav$_{15-133}$-PhoA
PstI
PhoA
EcoRV/MscI
138
Sav
NheI
15
Npr$_{ss}$-Sav$_{15-138}$-PhoA
PstI
PhoA
EcoRV/MscI
145
Sav
NheI
15
Npr$_{ss}$-Sav$_{15-145}$-PhoA
PstI
PhoA
EcoRV/MscI
159
Sav
NheI
15
Npr$_{ss}$-Sav$_{15-159}$-PhoA

FIG.7


EcoRV
KpnI
BamHI
pTZ18R
pBE311
lvs
pC194
Ap^r
Cm^r
HindIII
EcoRV
EcoRV
pBE653
sav
Ap^r
EcoRV
digestion &
ligation
EcoRV
EcoRV
BamHI
KpnI
pBE655
lvs-sav-lvs
Ap^r
Cm^r
HindIII

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 93/05240

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[6]

According to International Patent Classification (IPC) or to both National Classification and IPC

Int.Cl. 5 C12N15/31; C12N15/75; C12N15/62; C12N15/54

II. FIELDS SEARCHED

Minimum Documentation Searched[7]

Classification System	Classification Symbols
Int.Cl. 5	C07K ; C12N

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched[8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

Category[10]	Citation of Document,[11] with indication, where appropriate, of the relevant passages[12]	Relevant to Claim No.[13]
X	WO,A,9 203 560 (E.I. DU PONT DE NEMOURS AND COMPANY) 5 March 1992 see page 1, line 14 - line 32 see page 5, line 7 - page 6, line 6 see page 8, line 13 - page 11, line 3	1-8
X	WO,A,8 705 026 (THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK) 27 August 1987	15
A	see page 4, line 3 - page 5, line 10 see page 9, line 26 - page 10, line 7 see page 11, line 23 - page 13, line 25 & US,A,4 839 293 cited in the application	9,12,16, 18

-/--

° Special categories of cited documents :[10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
25 OCTOBER 1993	05 -11- 1993
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	MONTERO LOPEZ B.

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category °	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
A	WO,A,9 100 913 (E.I. DU PONT DE NEMOURS AND COMPANY) 24 January 1991 see page 3, line 15 - page 5, line 6 see page 8, line 10 - line 31 see page 10, line 13 - line 35	1,2,4-8

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO.

US 9305240
SA 75413

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information. 25/10/93

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9203560	05-03-92	AU-A-	8492291	17-03-92
		CA-A-	2067249	01-03-92
		EP-A-	0546049	16-06-93
WO-A-8705026	27-08-87	US-A-	4839293	13-06-89
		AU-A-	7165287	09-09-87
		EP-A-	0258411	09-03-88
		JP-T-	63502560	29-09-88
WO-A-9100913	24-01-91	US-A-	5162207	10-11-92
		CA-A-	2020598	08-01-91
		EP-A-	0483224	06-05-92
		JP-T-	4507045	10-12-92